Numerical Weather Prediction

Results of an initial stochastic physics scheme for the Met Office Unified Model

Forecasting Research Technical Report No. 452

Alberto Arribas

email: nwp_publications@metoffice.com

©Crown Copyright

Forecasting Research Technical Report 452
Results of an initial stochastic physics scheme for the Met Office Unified Model
A. Arribas

Abstract: **In the past decade, ensemble forecasting has proved to be a successful way of dealing with the uncertainty that limits the accuracy of weather and climate forecasts. However, most ensemble systems have generally paid more attention to the uncertainty in the initial conditions than to the model error, even thought it is now fully recognised that the representation of the uncertainty related to the use of imperfect models is unavoidable. Here we present a stochastic physics scheme developed for the Met Office Unified Model. This scheme consists of two independent components: a Random Parameters scheme in which a set of tunable parameters from different parameterizations are randomly perturbed thorough the simulation, and a stochastic convective vorticity scheme designed to account for the model error due to the failure of models to effectively represent the PV anomalies which are generated by Mesoscale convective systems.**

The impact of the stochastic scheme on the climatology of the model and its potential for short-range ensemble forecasting are analyzed. Our results show that the scheme is able to impact the forecast evolution on short time-scales with a neutral effect on the model's climatology.

History

Version 1.0	Revised by Ken Mylne	27 / 8 / 2004
Version 1.1	Circulated to Ken Mylne and Neill Bowler	8 / 9 / 2004
Version 1.2	Draft revised after final comments Approved by Ken Mylne	28 / 10 / 2004

Results of an initial stochastic physics scheme for the Met Office Unified Model

A. Arribas

1. - Introduction

In the past decade, ensemble forecasting has proved to be a successful way of dealing with the uncertainty that limits the accuracy of weather and climate forecasts. However, although generally more attention has been paid to the uncertainty in the initial conditions (Toth and Kalnay, 1993; Molteni et al., 1996) it is now fully recognised that the uncertainty related to the use of imperfect models also needs to be addressed in any ensemble system (Hou et al., 2001). In fact, the representation of model errors is thought to be an even greater challenge than simulating initial value related errors (Buizza et al., 2004).

In terms of model formulation, there are two main sources of uncertainty: first of all, only imperfect models are available, and second, the resolution of these models is limited. The place where both factors more clearly come together is in what is known as physical parameterizations (i.e. the representation of the effects of processes occurring at unresolved scales using comparatively simple deterministic functions of the resolved variables). In any of these parameterizations the value of a certain number of the empirical-adjustable parameters and thresholds is somewhat arbitrary, either because of being based on incomplete physical knowledge of the process or because of having been tuned to give optimal results for a test case that is not necessarily representative of more general applications (Yang and Arrit, 2002). Another downside of conventional or deterministic parameterizations is that, by returning something like a mean parameterized quantity, computed from the relevant large-scale variables, they neglect the unpredictable fluctuations of the small-scale processes which, in turn, could propagate upscale and force large-scale flow variability. Thus, even if unresolved scales only describe a small fraction of the total variance of the system, neglecting it can, in some circumstances, lead to gross forecast errors (Palmer, 2001). Plausibly, this is one of the reasons that explains the underdispersion (i.e. the observation lies outside the ensemble) often found in EPSs (Hamill and Colucci, 1997; Buizza et al., 2004).

Most of the research completed to date accounted for the model error uncertainties through the use of different model parameterizations or a multi-model approach. Just to mention some of the most recent studies, Houtekamer et al. (1996) used two different models and different physical parameterization schemes in the medium-range global ensemble run at the Meteorological Service of Canada; Hou et al. (2001) used 4 limited area models (LAM) and several combinations of changes in the model physical parameterizations of one of the LAMs [convection, moisture availability and boundary layer, as described in Stensrud et al. (1999)] for a short-range ensemble forecasting system, and Mylne et al. (2001) combined the ECMWF and Met Office models to show that, for medium-range, the multi-model approach outperformed each individual system. A different approach, using a range of plausible values for some parameters (but keeping those values fix in each ensemble member during the whole simulation), was followed by Yang and Arrit (2002) using a regional climate model and by Murphy et al. (2004) in the Quantifying Uncertainty in Model Prediction project (QUMP). In general, all these studies showed an improvement in the skill of the system (or a better estimation of the uncertainty associated with the forecast) when model error was accounted for. However, all multi-model systems are subject to the fundamental problem that they should consist of a series of model attractors that systematically bound the real-system attractor and, unfortunately, the real-system attractor is unknown, hindering their construction (Hansen, 2002).

A different approach is to develop explicit stochastic parameterizations which could have the advantage of allowing the forecast ensemble to explore important nearby regions of phase space that the deterministic parameterization would not reach. Interestingly, although the idea of stochastic parameterizations was suggested at least thirty years ago (Lorenz, 1975), few attempts have been made in this direction. The European Centre for Medium-Range Weather Forecast (ECMWF) pioneered this approach by including a stochastic perturbation to the net effect of parameterized physical processes in their ensemble prediction system (Buizza et al., 1999). Their results showed an improvement of the ensemble performance, particularly for the probabilistic prediction of precipitation. However, their scheme had a minimal impact on the spread of the ensemble, as seen in 500 hPa geopotential or mean sea level pressure (PMSL) fields. Lin

and Neelin (2002) introduced a stochastic component to the convection scheme by using a first-order autoregression model. Their stochastic scheme was designed to simulate nondeterministic processes that are not explicitly accounted for by the convection scheme and their results show that, by adding the stochastic scheme, the Global Circulation Model (GCM) simulates part of the low-frequency convective variance normally underestimated by GCMs. Bright and Mullen (2002) extended this approach to the Convection and Planetary Boundary Layer schemes of the fifth-generation Pennsylvania State University-National Center for Atmospheric Research Mesoscale Model (MM5; Grell et al., 1994) showing a slight increase in forecast skill. A more complex approach was developed by Gray and Shutts (2002; GS02 hereafter) in the Stochastic Convective Vorticity (SCV) scheme. The idea is to force a Potential Vorticity (PV) dipole, the scales of which are determined using a randomised function, associated with mesoscale convective systems onto the resolved scale of the GCM. Their results over a set of case studies showed that the scheme was able to produce realistic perturbations and preliminary tests within an ensemble showed an increase in the spread of the system (J. Denholm-Price, personal communication).

In this study we combine the SCV with a new Random Parameters scheme (RP) with the aim to create a more complete stochastic physics scheme for the Met Office Unified Model (UM) to be used in ensemble prediction systems. The combination of both schemes would allow us to better explore the up-scale propagation of uncertainty from the unresolved sub-grid scales and deficiencies of the parameterized processes. The paper is organised as follows: The stochastic physics scheme is described in section 2. Section 3 is dedicated to explaining the impacts of the scheme on a group of short-range forecast studies. Section 4 is focused on the impacts of the scheme on the climatology and inter-seasonal variability of the model. Conclusions are presented in section 5.

2. - System description

The stochastic physics (STPH) scheme employed in these experiments consists of the combination of two individual schemes: the Stochastic Convective Vorticity and the Random Parameters.

2.1. - The Stochastic Convective Vorticity (SCV) scheme

The main aim of the SCV scheme (GS02) is to represent a Potential Vorticity (PV) anomaly dipole similar to the one typically associated with a Mesoscale Convective System (MCS) in a GCM. As shown in observational studies (Bartels and Maddox, 1991), MCSs possess an upper level anticyclone, associated with the cirrus outflow or anvil, and also a smaller scale mid-level cyclonic vortex near the freezing level. However, given that MCSs may be only partially resolved in the model, the PV dipole generated by them may not be well represented. Consequently, the GCM may not transfer as much of the subgrid diabatic heating into resolved balanced motions as would be suggested from high resolution modelling studies (Gray et al., 1998). It has been also shown that PV anomalies associated with MCSs could have a significant impact on the forecast evolution, especially when located close to baroclinic zones (Gray, 2001; Beare et al, 2003).

In the SCV scheme the PV dipole is formed by two vortices, one mid-level cyclone representing the positive PV anomaly and one upper-level anticyclone representing the negative PV anomaly, the scales of which are determined using a randomised function. Because the application of a PV anomaly through a PV inverter would make the scheme prohibitively expensive, a simpler approach, just modelling the vorticity part of the PV anomaly, was used. Also, it appears from Gray (2001) that the vorticity is more important in perturbing the forecast than the stratification, so this simplification is not expected to be detrimental to the overall impact of the scheme.

Schematically, the upper-level anticyclone is of depth d_a, centred at height z_a and consisting of a zero PV core of radius a. The mid-level cyclonic vortex is of depth d_c, centred at height z_c and is of radius ηa, where η is set as 0.5. Therefore, we have six parameters (see Figure 1), a, d_a, d_c, z_a, z_c, and η that determines the size, shape, position and magnitude of the vorticity dipoles. In implementing the SCV all six could be given some form of random element but, given that the size of the MCS, a, is clearly the parameter which has the largest range in the atmosphere and in order to simplify the SCV scheme for initial implementation, only a will vary stochastically. As suggested by Shutts (1987), a is related to the value of CAPE (Convective Available Potential Energy) according to the following formula:

$$a = \alpha R p_r \frac{\sqrt{CAPE}}{f} \qquad (1)$$

Where α is a tuneable constant which was determined empirically to produce good agreement between simulated and observed frequency and size of MCSs. A value of 0.1 was found to produce satisfactory results. R is a random number and p_r is the diagnosed convective rain, which provides a link to the convection scheme in the model and makes a proportional to the magnitude of the diagnosed convective

rain rate. *f* is the Coriolis parameter. It must be noted that as $f \rightarrow 0$ at the Equator $a \rightarrow \infty$, to avoid this, *f* is assigned a minimum value of 2 x 10-5 s-1. A more complete description of the SCV scheme can be found in GS02.

2.2. - The Random Parameters (RP) scheme

There are many physical processes (convection, boundary layer exchanges, etc) that occur on scales too small to be directly resolved by the model and need to be parameterised. These parameterisations involve a number of empirical-adjustable parameters and thresholds which are given somewhat arbitrary values. The RP scheme aims to account for the uncertainty associated with these empirical parameters and to simulate the nondeterministic processes not explicitly accounted for by the different parameterizations.

The idea behind the RP, to treat a selected group of parameters as stochastic variables, is similar to the approach followed by Lin and Neelin (2000), and Bright and Mullen (2002). Thus, each parameter value is calculated using a first-order auto regression model (Wilks, 1995) as given by:

$$P_t = \mu + r(P_{t-1} - \mu) + \varepsilon \qquad (2)$$

Where P_t is the parameter value at time t, μ is the mean value of the parameter, *r* is the autocorrelation of *P* and ε is the stochastic shock term.

A total of 10 parameters from 4 different physical parameterizations are included in the 2004/05 version of the scheme (see Table 1). The physical parameterizations affected by the scheme are: Large scale precipitation (Wilson and Ballard, 1999), Convection (Gregory and Rowntree 1990), Boundary Layer (Lock et al. 2000) and Gravity Wave drag (Webster et al., 2003). A brief description of each parameter follows:

- *Entrainment Rate Coefficient*: Scalar, part of the convection scheme. This parameter scales the rate of mixing between environmental air and the convective plume.
- *Time scale for destruction of CAPE*: Scalar, part of the closure of the convection scheme. It controls the intensity of the convective mass flux.
- *Critical relative humidity*: Vector, part of the large scale precipitation scheme. This is the threshold of relative humidity for cloud formation. It takes different values on different model levels. In order to be coherent with the structure in the standard version of the model only the value at the upper levels (i.e. 10-38) is calculated using the first-order auto-regression model (Eq. 2). The value at the rest of levels are obtained by adding the following amounts to the values of the upper levels: 0.050 (levels 7-9), 0.100 (levels 3-6), 0.125 (level 2) and 0.150 (level 1).
- *Cloud droplet to rain conversion threshold*: Scalar, part of the large scale precipitation scheme. Its value is different for land and sea points.
- *Ice fall speed*: This is not properly a parameter in the UM version 6.0 so, in the RP scheme, this variable is modified through the CI parameter (see Table 1), which is involved in its calculation.
- *Flux profile parameter*: Scalar, part of the stable boundary layer scheme. It determines the stability dependence of turbulent mixing coefficients.
- *Asymptotic neutral mixing length*: Scalar, part of the boundary layer scheme. This parameter is required for calculation of turbulent mixing coefficients. This only affects the stable Boundary Layer mixing scheme.
- *Gravity Wave constant*: Scalar, part of the Gravity wave drag scheme. It defines the magnitude of the parameterized response.
- *Critical Froude number*: Scalar, part of the gravity wave drag scheme. It determines the proportion of the drag attributed to flow blocking and gravity wave drag respectively. The larger the critical Froude number the larger the proportion attributed to flow blocking.

In order to avoid unrealistic values, each parameter is bound by minimum and maximum values P_{max} and P_{min} which were estimated by experts in each field (see values in Table 1). The stochastic shock term, ε, is sampled from a uniform distribution from the range $\pm (P_{max} - P_{min})/3$. This is an arbitrary choice, but a uniform distribution was preferred over a normal distribution to avoid ε values being constrained towards a mean value. The autocorrelation *r* has been given a value of 0.95 and the parameters' values are updated every 3-hours. It must be recognised that, because no knowledge exists a priori of ε and *r* values, both were determined empirically after some limited 'tuning' of the RP scheme. Finally, the parameters' values are equal for all grid points (i.e. the spatial correlation is 1).

3. - Analysis of the short-range experiments

The GCM used in these experiments is the Met Office Unified Model (UM) version 6.0, run at a resolution of 0.83º longitude by 0.55º latitude and 38 vertical levels (the operational resolution and configuration as in April 2004). Given that the STPH scheme was originally conceived as part of a short-range EPS, the first test to be applied in evaluating the scheme is to determine whether it has any impact within a GCM in this time-scale. Three case studies - chosen as episodes of intense rain took place in the European region - were run using analysis at 00Z from each selected case: 20040127; 20040220 and 20040326. Four simulations were run for each case: a control run using the standard version of the model and a 3-member ensemble in which each member includes the STPH scheme with different random seeds. Each forecast was run for 72 hours.

Table 2 shows the root mean square (rms) differences between the STPH runs and the control run for pressure at mean sea level (PMSL). As can be seen, the impact of the STPH scheme varies, not only from case to case but for each member in a particular case. The largest differences after 72-h are found in the February case, with typical values of 0.5 hPa. In general, these rms differences at T+72 are similar to those found by GS02 at T+120 when testing the SCV scheme, which implies that the RP scheme is having a positive contribution, increasing the spread of the ensemble of stochastic runs in the first 72-h. This is further confirmed by the fact that GS02 found the largest differences produced by the SCV scheme confined over the southern hemisphere storm track, with little impact in other regions of the globe, whereas, as shown in Figure 2 (corresponding to the 20040326 case at T+72, but similar results were obtained for the other two cases), the impacts of the STPH scheme extend to the northern hemisphere and some areas of the tropics.

One of our main interests is the impact of the STPH scheme on the surface variables. Figure 3 shows the differences between the 3-member ensemble mean and the control simulation for 2 meter temperature and total accumulated precipitation for the 20040220 case. Differences in temperature are noticeable, reaching values over 5 degrees in some regions. Not surprisingly, given that all simulations use the same prescribed sea surface temperatures (SSTs), temperature differences are mainly concentrated over land. In general, these differences are quite scattered but there are some areas where differences are extended over larger regions (e.g. North America, Siberia). Differences in accumulated total precipitation (including rain and snow) between the ensemble mean and the control simulation, along with contours showing the control precipitation over 10 mm, are shown in Figure 3b. Most of the differences are found where the control run is already forecasting some precipitation. In other words, the STPH scheme is mainly altering the magnitude of the forecast precipitation - to quite a large extent, up to 30 mm or 40% of the control value, but not significantly changing the geographical areas affected. This is a common characteristic of all case studies. However, this does not imply that the scheme results are not useful. For example, an intense snowfall happened on the 28th and 29th of January 2004 over the United Kingdom that was poorly forecast but, if we look at the STPH ensemble members (Figure 4), one of them show an intense drop in surface temperature which, in conjunction with an intensification of the front that produced most of the snowfall, could have been a helpful sign for the forecasters.

The STPH scheme is, therefore, having a substantial impact over the surface variables and it seems to positively increase the spread. An analysis of the contribution of each individual scheme (SCV and RP) to the spread of the system was analyzed by running 6 additional forecasts for each case: 3 of them only using the SCV scheme and 3 of them with only the RP scheme. Results for the 20040326 are shown in Figure 5. From this, it is clear that the increase in the spread is non-linear; despite that the spread of the RP scheme is almost double that of the SCV scheme, the total spread of the STPH system is smaller than the sum of the individual contributions. This suggests that both schemes could be exploring the same regions of phase space and that the spread of the system could be linked to the overall synoptic evolution of the forecast rather than necessarily the random numbers. Figure 6 shows the evolution of the (zonal mean) spread for all case studies. As can be seen in these hovmoller diagrams, the spread varies from case to case and also geographically. Thus, in the January case the largest spread after 72-h is found around 55N and 15S; in the February case the largest spread is found in both hemispheres at high-latitudes (55-70N and 80S), whereas the March case shows a more evenly distributed spread, with maxima concentrated on the mid-high latitudes (60-70N) on the Northern Hemisphere, equatorial regions and high latitudes (70-80-S) on the Southern Hemisphere. In all three cases, the spread is generated almost uniformly over the whole domain when the STPH scheme kicks off, but it quickly concentrates and increases only in certain regions, supporting the idea that the spread generated by the STPH scheme is linked to the background flow.

A second experiment (using 12Z analysis from 20040129) was performed to analyze the growth of the STPH perturbations and compare it with the behaviour of perturbations to the initial conditions. Now we ran three 8-member ensembles: in the first one, called IC, only the initial conditions were perturbed. The second one, called ME (from Model Error) included the STPH scheme but no perturbations to the initial conditions. The third one included both, perturbations to the initial conditions and the stochastic physics scheme (IC+ME). The perturbations to the initial conditions were calculated through an Ensemble Transform Kalman Filter approach (Bishop et al., 2001). The spatial distribution of PMSL spread after 72 hours is shown in Figure 7

for the three ensembles: IC+ME (above), IC (center) and ME (bottom). It can be seen that in all three ensembles the larger amount of spread is found over the same regions: North Atlantic-Scandinavia, east of the Japanese coast and near the North American west coast. So, same as we noted previously, it seems that the spatial distribution of spread does not depend critically on the mechanism used to introduce the perturbations. This implies that the stochastic physics scheme and the initial conditions will be exploring common areas of the phase space and the increase in spread will be non-linear. Thus, if we compare the spread in the IC and ME ensembles, it seems that the ME spread is approximately a quarter of the IC spread. However, when perturbations to the initial conditions and the stochastic physics scheme are both part of the ensemble, as in IC+ME, the amount of spread is increased a mere 5% respect to the IC ensemble. This can be more clearly observed in Figure 8, where the time evolution of the average spread over the northern hemisphere is shown.

Based on these experiments, the role of the stochastic physics scheme seems to fit well with the Tribbia and Baumhefner (2004) view of the inverse cascade of energy from smaller to larger scales. According to their view, the primary role of the inverse cascade is to seed disturbance in the baroclinically active region. Once in the synoptic scale, the perturbations will organize within the synoptic structures and amplify, extracting energy from the large-scale background flow, not the smaller scales. This effect can be observed in the analysis of the kinetic energy spectra of the forecasts.

Figure 9 shows the kinetic energy spectra at T+72 for the control simulation and one of the STPH runs for the 20040326 case. As can be seen, the spectra are almost identical, with energy levels that are generally of the correct magnitude (with a slope near -3) between approximately 4000 and 1000 km wavelength, but failing to reproduce the -5/3 spectra typically observed in the mesoscale (Nastrom et al., 1984). This difference from the observed spectra in the mesoscale is due to two factors, a lack of resolution of the model and an excessive model dissipation, which removes energy and depresses the spectra. The horizontal resolution used for these experiments is N216, roughly equivalent to 60 km but, if we use the spectra to define the effective resolution of a model as the wavelength where a model's spectrum begins to decay relative to the observed spectrum (Skamarock, 2004), we can see that the effective resolution is around 450 km which partially explains the inability of the model to reproduce the -5/3 spectra of the mesoscale. As mentioned, the STPH spectrum is almost identical to the spectrum of the control run. On the positive side, this means that the STPH scheme is not leading to an unrealistic build up of energy. However, it also implies that the scheme has not prevented the excessive model dissipation seen in the largest wave numbers (i.e. smaller scales) and, therefore, not succeeded to up-scale energy from the sub-grid scales to explicitly resolved scales.

A similar message arises when the spectra of the IC and ME ensembles are analyzed. Figure 10 shows the spectra of 500 geopotential height (averaged from the 8 ensemble members) at (a) T+12 and (b) T+72. Again, the spectra of both ensembles are very similar, although interestingly, whereas the ensemble-average have filtered most of the noise in the IC spectra, the ME spectra is not so smooth. In fact, it seems that the noise upscale to larger wavelengths form T+12 to T+72. This could be indicative of inertia-gravity modes continuously generated by the STPH scheme. Also at T+72, it is clearer an upturn of the spectral tail at higher wave numbers in the ME spectra but not to an extent that indicates a build up of energy to physically unrealistic levels (Skamarock, 2004).

4. - Analysis of the climate experiments

As shown in the previous section, the STPH scheme does have an impact within the UM on the short-range. The next step is to determine the effect it has on the UM's climatology, and whether it is beneficial, detrimental or neutral. A climate simulation can be also useful to determine the impact of the scheme on a seasonal timescale.

Two 3-year simulations, a control run using the standard version of the model and a run including the STPH scheme, were completed using HadGAM1, the climate version of the Met Office Unified Model. This is an atmosphere-only model evolved from HadAM3 (the model used operationally for climate research at the Hadley Centre until 2004, Pope et al., 2000) with a horizontal resolution of 3.75 x 2.5 degrees (lon-lat) and 38 levels in the vertical. SSTs were prescribed using observed data.

Although three years is too short for a full quantitative analysis of climate differences between the two runs, it is long enough to spot any significant systematic changes that the presence of the stochastic scheme may cause. Both simulations were evaluated using the Climate Prediction Index (CPI; Sexton et al., 2002), which was developed to measure how well a climate model reproduces various aspects of the climate system such as atmospheric radiation and clouds, atmospheric dynamics, the hydrological cycle and surface fluxes. A complete list of all variables used can be found on the x-axis in Figure 11. For all simulated seasonal means

(DJF, MAM, JJA, SON), each variable is compared against the appropriate observational or reanalysis data set, using a normalised version of the root-mean square error. This statistic penalises bias, differences in the spatial variances of the observed and modelled means and poor pattern correlation. A perfect run would get a CPI value equal to 0.

The CPI values for both simulations are similar, 2.461 for the control run and 2.534 for the stochastic physics run. These values can be considered very good (although it depends on the exact model configuration, typical values for the previous version of the model, HadAM3, were around 3.0). The total CPI value for the stochastic run is slightly poorer than the control run. This is mainly due to the contribution from the relative and specific humidity components of the CPI (see Figure 11), and it is caused by the impact of the stochastic scheme on the "critical relative humidity" parameter. It seems that the minimum value for this variable inside the RP scheme may be too low, something that could be tuned in future versions of the scheme to produce better results. However, there are other variables in which the stochastic run is outperforming the control run, especially in the "dynamical" variables, such as 250mb velocity potential or the different stream function components. Therefore, we can conclude that both simulations are providing similar quality representations of the current climate and that the stochastic scheme is not having a strong impact (either positive or negative) on the model climate. This is very encouraging, especially considering that we are comparing the stochastic scheme against a well tuned control version of the model.

Having ascertained that the stochastic scheme is having a neutral impact on the model's climate, it is interesting to assess whether it is capable of affecting the forecast evolution on seasonal time-scales. Figure 12 shows the PMSL fields for DJF at year 3. The top panel corresponds to the control run, the middle one to the stochastic physics run, and the differences between both are shown in the bottom panel. As can be seen, differences are substantial. In fact, each run is simulating a different phase of the North Atlantic Oscillation (NAO). Thus, whereas the control run is simulating a positive NAO phase, with stronger winds crossing the Atlantic Ocean on a more northerly track, the run including the stochastic physics scheme is simulating a negative NAO phase, with winds following a more west-east pathway. As shown in Figure 13, this is clearly reflected in the seasonal precipitation patterns, producing a decrease of precipitation over a region centered on the Iberian Peninsula and an increase over Scandinavia. Therefore, even when the initial and boundary conditions (SSTs) are kept unchanged, the inclusion of sub-grid scale variability may have a substantial impact on the large-scale flow (i.e. low frequency variability) characteristics on seasonal time-scales.

5. - Summary and conclusions

A new stochastic physics scheme (STPH) designed to represent the uncertainty due to model error in the Met Office Unified Model is presented in this paper. The STPH scheme is formed by two different components: the stochastic convective vorticity and the random parameters schemes, which are intended to represent the uncertainties arising from the unresolved sub-grid scales and deficiencies of the parameterized processes respectively. Preliminary results from this scheme are encouraging. The stochastic physics scheme is capable of producing a substantial impact on surface variables on time-scales varying from short-range to seasonal while maintaining a neutral effect on the model's climatology.

The performance of the system in the short-range (up to 72-h) was analyzed using three case studies. Four simulations (a control run and a 3-member ensemble in which each member includes the STPH with a different random seed) were completed for each case study. The local impact of the stochastic physics scheme on the surface variables is noticeable: up to 5ºC in 2 m temperature, 50 mm/d in precipitation and 5 mb in PMSL. When compared with previous studies using only the SCV scheme, the new STPH scheme also shows a better performance in terms of spread (larger values are found during the first 72-h) and a broader geographical impact. The impact of the SCV-only scheme was limited to the Southern Hemisphere storm track, whereas the new STPH is also affecting areas on the Northern Hemisphere and the tropics.

The results show that the spread growth is not linear. Thus, the spread of the STPH is smaller than the individual contributions from the RP and SCV schemes (the larger contribution comes from the RP scheme almost doubling that of the SCV) suggesting that both schemes are exploring the same regions of phase space. This was confirmed by a second set of simulations in which the spread of different ensembles - in which the perturbations were applied only to the initial conditions (IC), only to the model errors (ME) or both (IC+ME) - is compared. This comparison shows some interesting characteristics: on the one hand, when IC and ME are compared, the spread in ME is approximately a quarter of the spread due to the ICs. However, when both perturbations (IC+ME) are included in the ensemble, the increase in the spread with respect to the IC-only ensemble is much smaller (~5 %). Also, after 72 hours all three ensembles (IC, ME and IC+ME) produce their largest spread over the same regions. This indicates that the growth and spatial distribution of spread is not exclusively related to the mechanism used to introduce the perturbations, but to the

background flow. The STPH works in a similar way to the alternative view of the inverse cascade of energy proposed by Tribbia and Baumhefner (2004), it seeds disturbances in the mesoscale region where, in the baroclinically active areas, they amplify by extracting energy from the large-scale background flow, not the smaller scales. The analysis of the energy spectra (with little differences between the STPH and the control runs) supports this hypothesis.

The impact of the STPH scheme on the model's climatology was also analyzed. For this purpose, two 3-year long simulations (using the standard version of the model with or without the stochastic physics scheme) were completed. Both simulations were evaluated using the Climate Prediction Index which measures how well a climate model reproduces various aspects of the climate system. The CPI values for both simulations are similar and quite good, showing that both runs provide similar quality representations of the current climate and that the STPH is having a neutral impact on the model's climate. This is encouraging, especially considering that we are comparing the stochastic scheme against a well tuned control version of the model. However, despite having a neutral impact on the model's climatology, the STPH has been found able to impact large-scale flow characteristics, such as the NAO, on seasonal time-scales.

The fact that the impacts found in very different time-scales are non-negligible confirms that an explicit representation of model error is needed in ensemble prediction systems, independently of the time-range they were originally designed for.

References

Bartels, D.L. and Maddox, R.A., 1991: Midlevel cyclonic vortices generated by mesoscale convective systems. Mon. Wea. Rev., 119, 104-118.

Beare, R.J., Thorpe, A.J. and White, A.A., 2003: The predictability of cyclones: non-linear sensitivity to localised potential vorticity perturbations. Quart. J. Roy. Meteorol. Soc., 129, 219-237.

Bishop, C.H., Etherton, B.J. and Majumdar, S.J., 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. Mon. Wea. Rev., 129, 420-436.

Buizza, R., Miller, M. and Palmer T.N., 1999: Stochastic representation of model uncertainties in the ECMWF EPS. Q.J.R. Meteorol. Soc., **125**, 2887-2908.

Buizza, T., Houtekamer, P.L., Toth, Z., Pellerin, G., Wei, M., Zhu, Y., 2004: Assessment of the status of global ensemble prediction. Submitted to Mon. Wea. Rev.

Bright, D.R. and Mullen, S.L., 2002: Short-range ensemble forecasts of precipitation during the southwest monsoon. Wea. Forecasting, **17**, 1080-1100.

Gray, M.E.B., Shutts, G.J. and Craig, G.C., 1998: The role of mass transfer in describing the dynamics of mesoscale convective systems. Quart. J. Roy. Meteorol. Soc., **124**, 1183-1207.

Gray, M.E.B., 2001: The impact of mesoscale convective system potential vorticity anomalies on numerical weather prediction forecasts. Quart. J. Roy. Meteorol. Soc., **127**, 73-88.

Gray, M.E.B. and Shutts, G.J., 2002: A stochastic scheme for representing convectively generated vorticity sources in general circulation models. APR Turbulence and Diffusion Note No. 285. Met Office, UK.

Gregory, D., and Rowntree, P.R., 1990: A mass flux convection scheme with representation of cloud ensemble characteristics and stability-dependent closure. Mon. Wea. Rev., **118**, 1483-1506.

Grell, G.A., Dudhia, J. and Stauffer, D.R., 1994: A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech. Note NCAR/TN-398+STR, 121 pp. [Available from MMM Division, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307]

Hansen, J.A., 2002: Accounting for Model Error in Ensemble-based state estimation and forecasting. Mon. Wea. Rev., **130**, 2373-2391.

Hamill, T.M. and Colucci, S.J., 1997: Verification of Eta-RSM short-range ensemble forecasts, Mon. Wea. Rev., **125**, 1312-1327.

Hou, D., Kalnay, E. and Droegemeier, K.K., 2001: Objective verification of the SAMEX'98 ensemble forecasts. Mon. Wea. Rev., **129**, 73-91

Houtekamer, P.L., Lefaivre, L., Derome, J., Ritchie, H. and Mitchell, H.L. 1996: A system simulation approach to ensemble prediction. Mon. Wea. Rev., **124**, 1225-1242.

Lin, J. W.-B., and J.D. Neelin, 2000: Influence of a stochastic moist convective parameterization on tropical climate variability. Geophys. Res. Lett., **27**, 3691-3694.

Lock, A.P., Brown, A.R., Bush M.R., Martin, G.M., and Smith, R.N.B, 2000: A new boundary layer mixing scheme. Part I: Scheme description and single column model tests. Mon. Wea. Rev., **128**, 3187-3199.

Lorenz, E.-N., 1975: Climate predictability. The physical basis of climate modelling, W.M.O. GARP Pub.Ser., 16, 132-136.

Molteni, F., Buizza, R., Palmer, T.N. and Petroliagis, T. 1996: The ECMWF Ensemble Prediction System: Methodology and validation. Q.J.R. Metorol. Soc., **122**, 73-119.

Murphy J.M., Sexton, D.H., Barnett, D.N., Jones, G.S., Webb, M.J., Collins, M. and Stainforth, D.A., 2004: Quantification of modelling uncertainties in a largae ensemble of climate change simulations. *Nature*, **430**, 768-772.

Mylne, K.R., Evans, R.E., Clark, R.T., 2001: Multi-model multi-analysis ensembles in quasi-operational medium-range forecasting. Q.J.R. Metorol. Soc., **128**, 361-384

Nastrom, G.D., Gage, K.S. and Jasperson, W.H., 1984: The atmospheric kinetic energy spectrum, 10^0-10^4 km. Nature, **310**, 36-38.

Palmer, T.N., 2001: A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parametrization in weather and climate pres-diction models. Q.J.R. Metorol. Soc., **127**, 279-304.

Pope, V.D., Gallani, M., Rowntree, P.R. and Stratton, R.A., 2000: The impact of new physical parameterisations in the Hadley Centre climate model HadAM3. *Clim. Dyn.*, **16**, 123-146.

Sexton, D.M., Barnett, D.N., Murphy, J., Jones, G.S. and Webb, M., 2002: Measuring the quality of HadSM3 simulations and its impact on the uncertainty of climate sensitivity. Hadley Centre report for DEFRA 14/3/02 [Available from Met Office, Exeter, EX1 3PB, UK]

Stensrud, D.J., H.E Brooks, J. Du, M.S. Tracton and E. Rogers, 1999: Using ensembles for short-range forecasting. Mon. Wea. Rev., **127**, 433-446.

Shutts, G. J., 1987: Balanced flow states resulting from penetrative, slantwise convection. J. Atmos. Sci., **44**, 3363-3376.

Skamarock, W.C., 2004: Evaluating mesoscale NWP models using kinetic energy spectra. Mon. Wea. Rev., in press.

Toth, Z. and Kalnay, E., 1993: Ensemble Forecasting at the NMC: The generation of perturbations. Bull. Amer. Meteor. Soc., **74**, 2317-2330.

Tribbia, J.J. and Baumhefner, D.P., 2004: Scale interactions and atmospheric predictability: An updated perspective. Mon. Wea. Rev., **132**, 703-713.

Webster, S., Brown, A.R., Cameron, D.R. and Jones, C.P., 2003. Improvements to the representation of orography in the Met Office Unified Model. Q. J. R. Meteorol. Soc., 129, 1989-2010

Wilks, D.S., 1995: Statistical methods in the atmospheric sciences – an introduction. International Geophysics Series. Vol **59**, Academic Press, New York, USA

Wilson, D.R. and Ballard, S.P., 1999. A microphysically based precipitation scheme for the UK Meteorological Office Unified Model. Q. J. R. Meteorol. Soc., 125, 1607-1636

Yang, Z. and Arrit, R.W., 2002: Tests of a perturbed physics ensemble approach for regional climate modelling. J. Clim., 15, 2881-2896.

Figure Captions

Figure 1. An example of a vorticity dipole as generated by the SCV scheme. Contours are of the component of velocity perpendicular to the vertical plane shown, with positive (negative) values shown by solid (dashed) contours. Figure reproduce with permission from Gray and Shutts (2002).

Figure 2. PMSL impacts of the STPH scheme on the 20040326 case at T+72. The top-left panel shows the ensemble standard deviation (shaded colours) and PMSL control values (contour). The other three panels show the differences between each ensemble member and the control run.

Figure 3. Differences between the 3-member ensemble mean and the control simulation for 2 meter temperature (left) and total accumulated precipitation in 72-h (right) for the 20040220 case.

Figure 4. Differences between one of the ensemble members and the control simulation for 2 meter temperature (left) and accumulated snow in 72-h (right) for the 20040127 case.

Figure 5. Time-evolution of PMSL spread for three configurations of the STPH scheme: only using the SCV scheme (dash-dot line), only using the RP scheme (dash line), combining both RP and SCV schemes (solid line). The results correspond to the 20040326 case.

Figure 6. Hovmoller diagrams showing the time-evolution of the zonal mean spread for PMSL and all case studies: 20040127 (top), 20040220 (center) and 20040326 (bottom).

Figure 7. Spatial distribution of spread for PMSL after 72-h (shaded) and ensemble mean values (contour). Three ensemble configurations are show: perturbations to initial conditions and model error (IC+ME, top), only perturbing the initial conditions (IC, center) and only model error (ME, bottom).

Figure 8. Time-evolution of PMSL spread for three different ensemble configurations: IC+ME (dash line), IC (solid line) and ME (dash line).

Figure 9. Kinetic energy spectra at 500 hPa and T+72 for the control simulation (black line) and one of the runs with the STPH scheme for the 20040326 case.

Figure 10. Kinetic energy spectra at 500 hPa for the IC (only perturbing the initial conditions) and ME (only including the STPH scheme as representation of the model error) ensembles at **(a)** T+12 and **(b)** T+72. The plotted energy spectra corresponds to the average from the 8 ensemble members.

Figure 11. Values of the Climate Prediction Index (CPI) and its components. The 24 components represent surface and atmospheric variables, and are calculated as the r.m.s. difference between simulated and observed present-day climatological mean patterns divided by the r.m.s. value of the standard deviation of simulated interannual variations. The CPI is calculated as the r.m.s. value of the 24 components. All components are weighted equally, apart from the fields of cloud cover which receive a relative weight of 1/3 since observations of high, medium and low cloud are interdependent for a given thickness category. OLR, outgoing longwave radiation; OSW, outgoing shortwave radiation; LW, longwave radiation; SW, shortwave radiation; TOA, top of atmosphere; MSLP, mean sea level pressure.

Figure 12. Mean sea level pressure for the winter period (December, January, February) at year 3. The control run is shown in the top panel, the run including the STPH scheme is shown in the central panel and the differences between both are shown in the bottom panel.

Figure 13. Differences in precipitation (mm/d) for the winter period (December, January, February) at year 3 between the control run and the run including the STPH scheme.

Parameter	Scheme	Min / std / Max values
Entrainment rate coefficient (*entcoef*)	Convection	2 / 3 / 5
Cape timescale (*cape_timescale*)	Convection	30 / 30 / 120
Critical relative humidity (*rhcrit*)	Large scale cloud	0.6 / 0.8 / 0.9
Cloud to rain conversion threshold over land (*cw_land*)	Large scale cloud	1E-4 / 8E-4 / 1E-3
Cloud to rain conversion threshold over sea (*cw_sea*)	Large scale cloud	5E-5 / 2E-4 / 5E-4
Parameter controlling ice fall speed (*CI*)	Large scale cloud	17 / 25.2 / 33
Flux profile parameter (*par_mezcla*)	Boundary layer	5 / 10 / 20
Neutral mixing length (*G0*)	Boundary layer	0.05 / 0.15 / 0.5
Gravity wave constant (*kay_gwave*)	Gravity wave drag	7.5E-4 / 7E-4 / 1.3E-4
Froude number (*gwd_frc*)	Gravity wave drag	2 / 2 / 4

Table 1. Parameters included in the RP scheme. The physical parameterization they are part of is indicated in the second column. The third column indicates the minimum and maximum limits for the value of the parameter together with the parameter's value as in the standard setting of the Met Office Unified Model version 6.0.

	2004 01 27			**2004 02 20**			**2004 03 26**		
	Run 1	Run 2	Run 3	Run 1	Run 2	Run 3	Run 1	Run 2	Run 3
24 – h	0.15	0.17	0.18	0.19	0.18	0.20	0.22	0.20	0.19
48 – h	0.19	0.24	0.26	0.30	0.35	0.27	0.31	0.30	0.26
72 – h	0.30	0.37	0.33	0.53	0.50	0.49	0.42	0.47	0.41

Table 2. r.m.s. differences between the control run and each of the runs including the STPH scheme for mean sea level pressure and all three case studies.

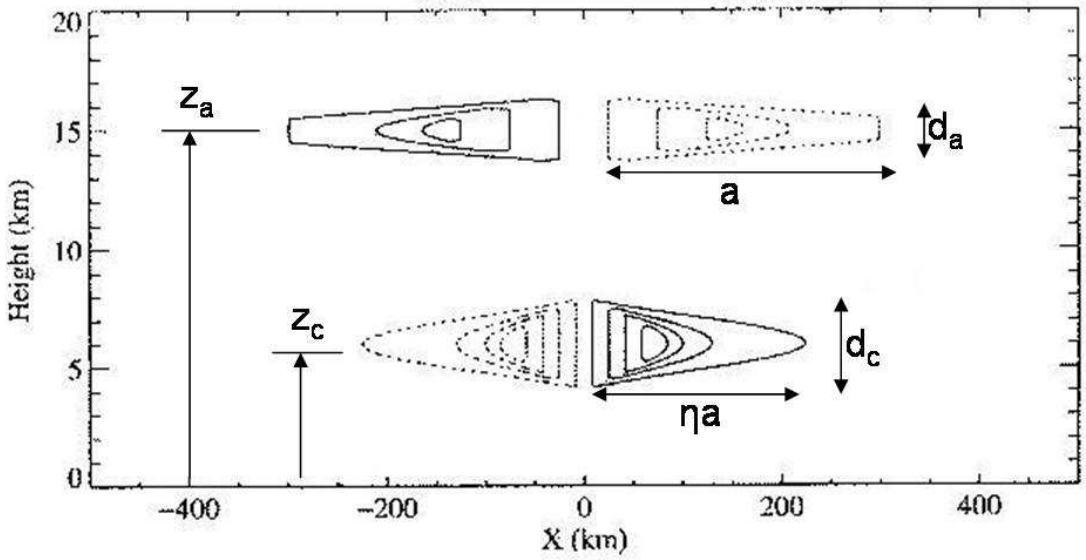

Figure 1. An example of a vorticity dipole as generated by the SCV scheme. Contours are of the component of velocity perpendicular to the vertical plane shown, with positive (negative) values shown by solid (dashed) contours. Figure reproduce with permission from Gray and Shutts (2002).

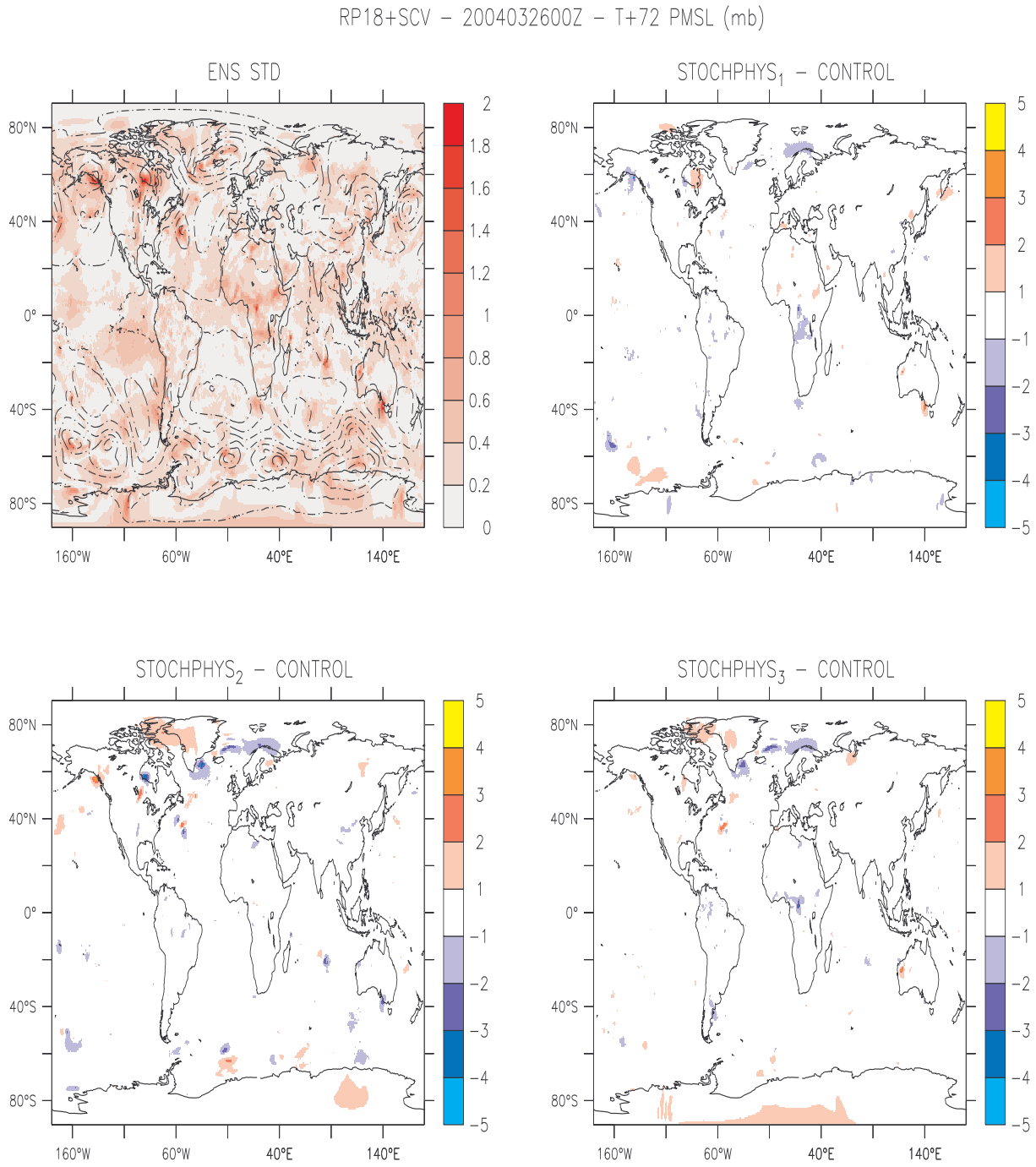

Figure 2. PMSL impacts of the STPH scheme on the 20040326 case at T+72. The top-left panel shows the ensemble standard deviation (shaded colours) and PMSL control values (contour). The other three panels show the differences between each ensemble member and the control run.

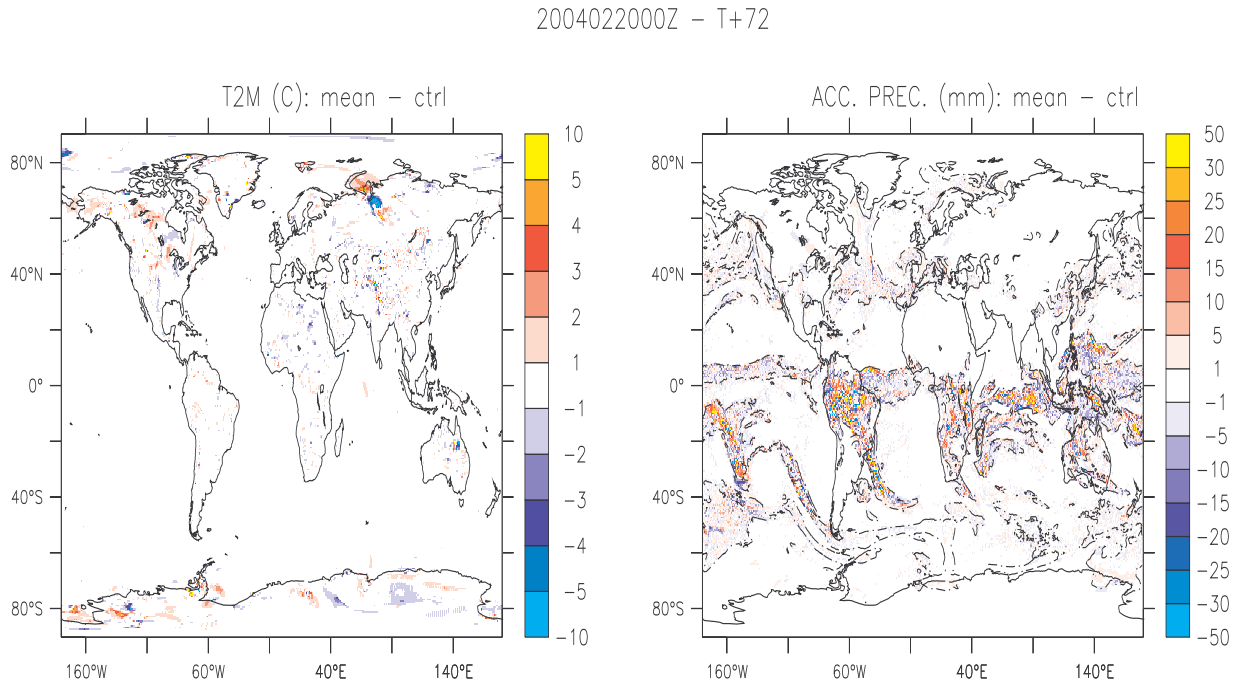

Figure 3. Differences between the 3-member ensemble mean and the control simulation for 2 meter temperature (left) and total accumulated precipitation in 72-h (right) for the 20040220 case.

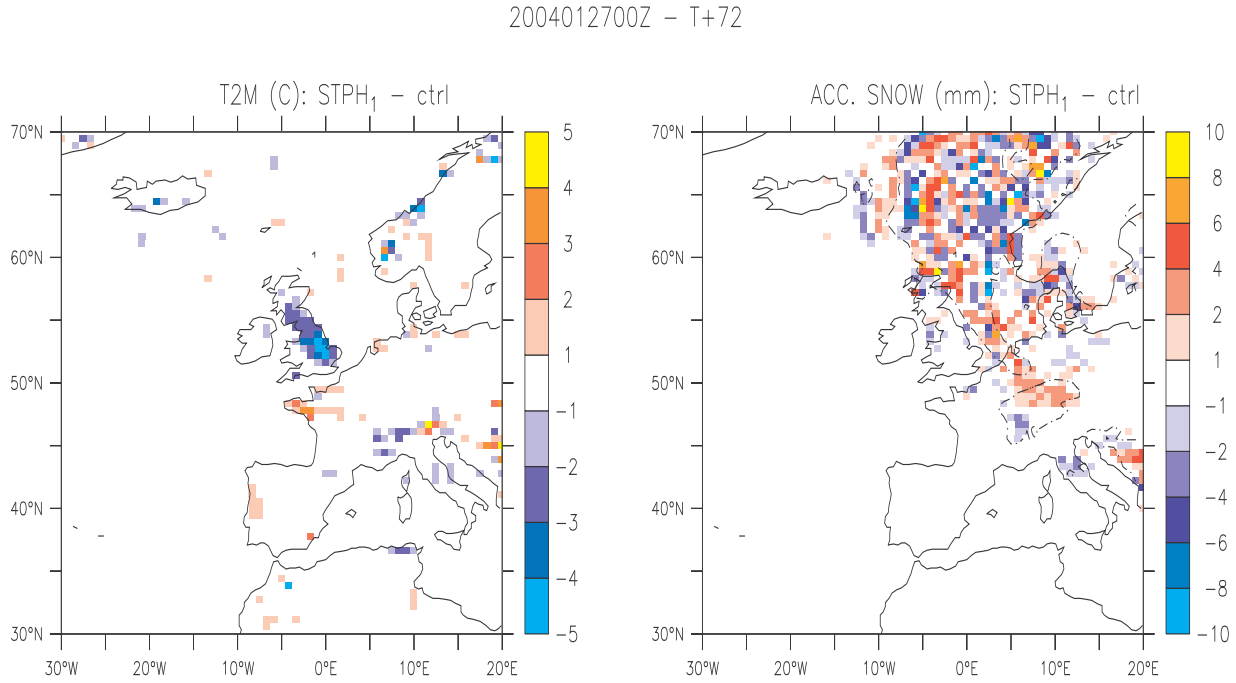

Figure 4. Differences between one of the ensemble members and the control simulation for 2 meter temperature (left) and accumulated snow in 72-h (right) for the 20040127 case.

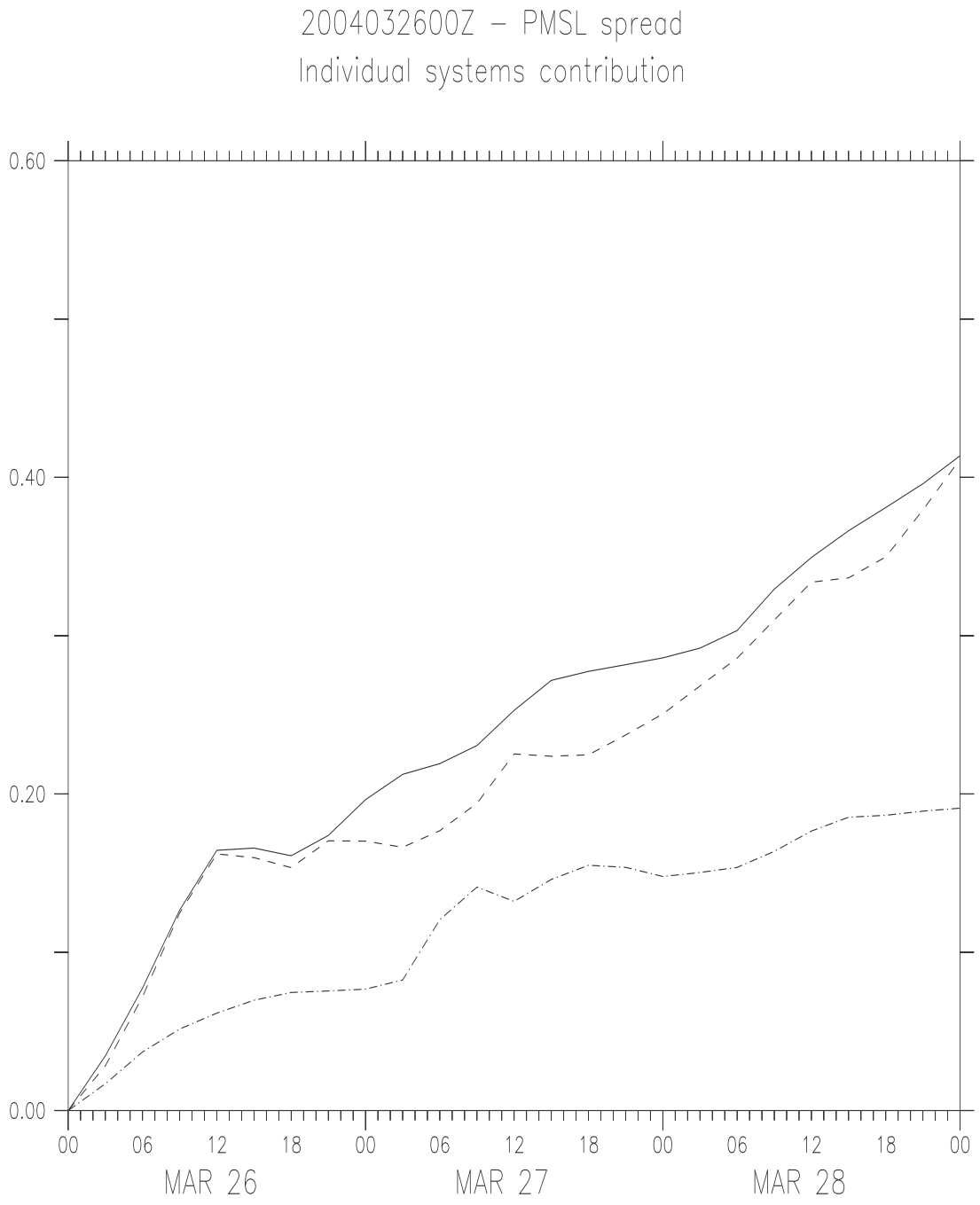

Figure 5. Time-evolution of PMSL spread for three configurations of the STPH scheme: only using the SCV scheme (dash-dot line), only using the RP scheme (dash line), combining both RP and SCV schemes (solid line). The results correspond to the 20040326 case.

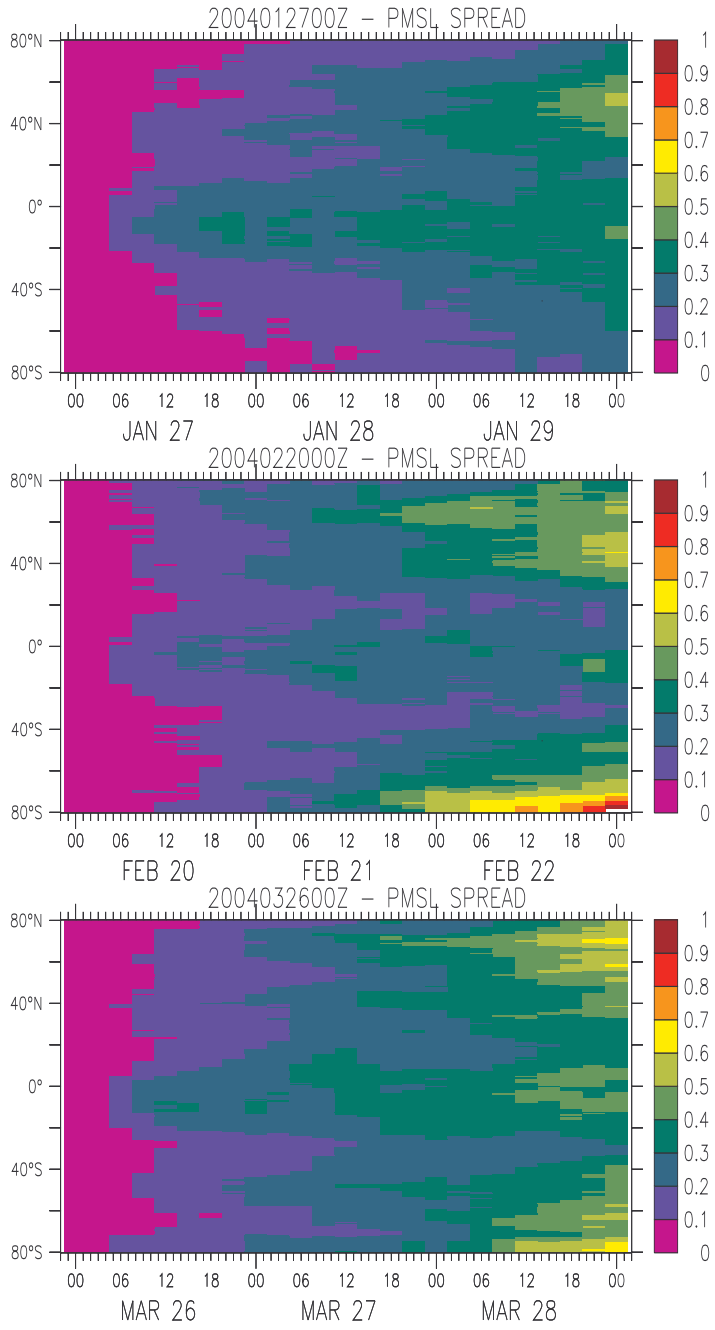

Figure 6. Hovmoller diagrams showing the time-evolution of the zonal mean spread for PMSL and all case studies: 20040127 (top), 20040220 (center) and 20040326 (bottom).

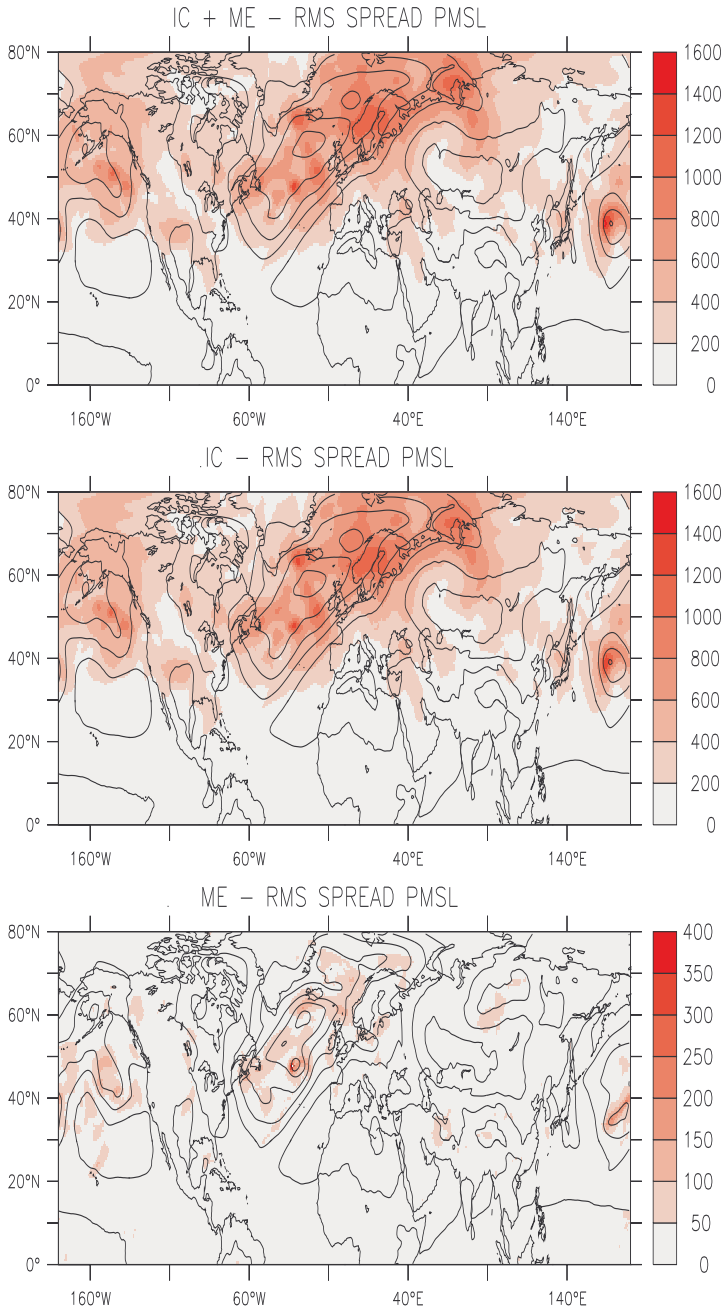

Figure 7. Spatial distribution of spread for PMSL after 72-h (shaded) and ensemble mean values (contour). Three ensemble configurations are show: perturbations to initial conditions and model error (IC+ME, top), only perturbing the initial conditions (IC, center) and only model error (ME, bottom).

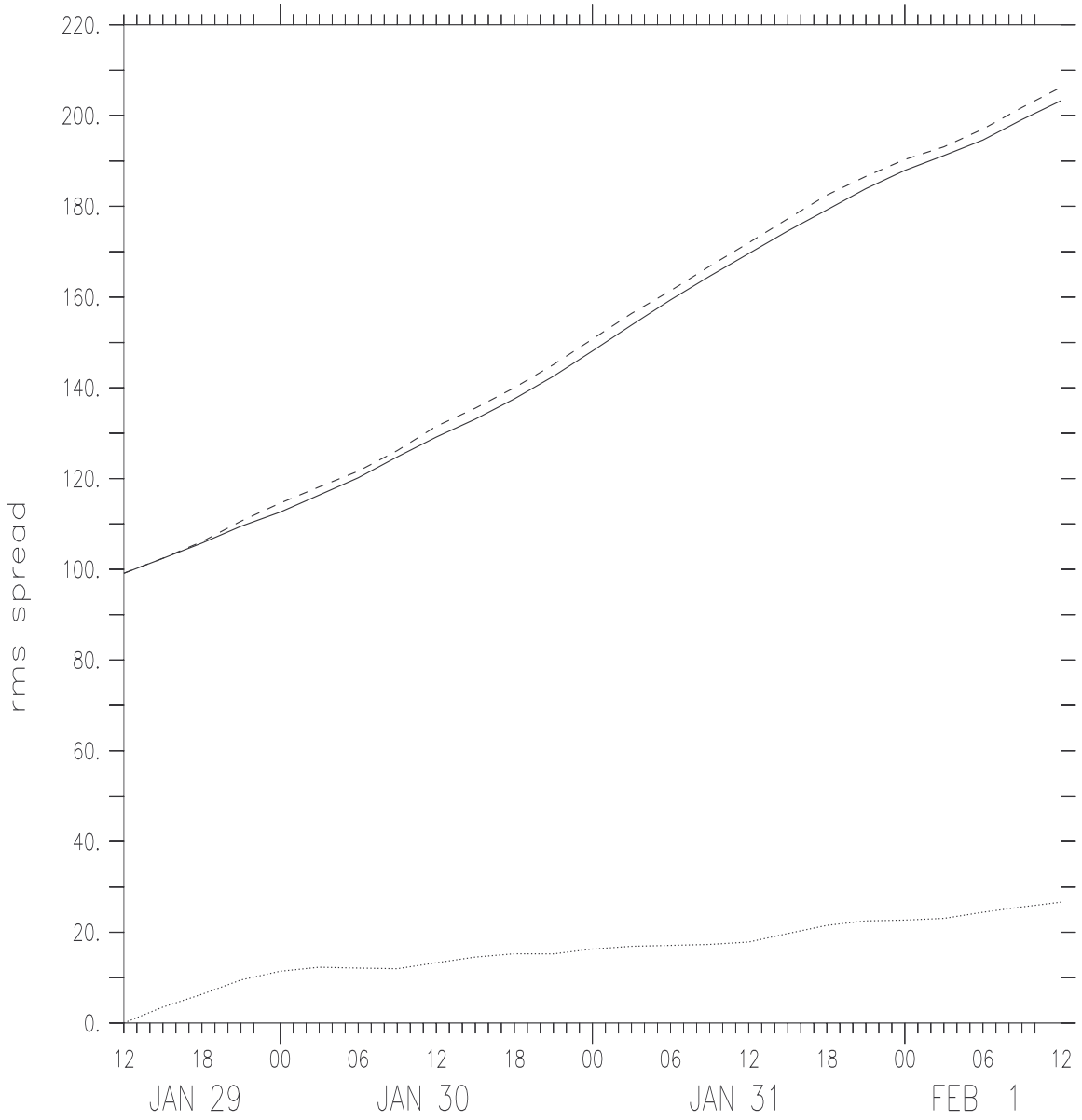

Figure 8. Time-evolution of PMSL spread for three different ensemble configurations: IC+ME (dash line), IC (solid line) and ME (dash line).

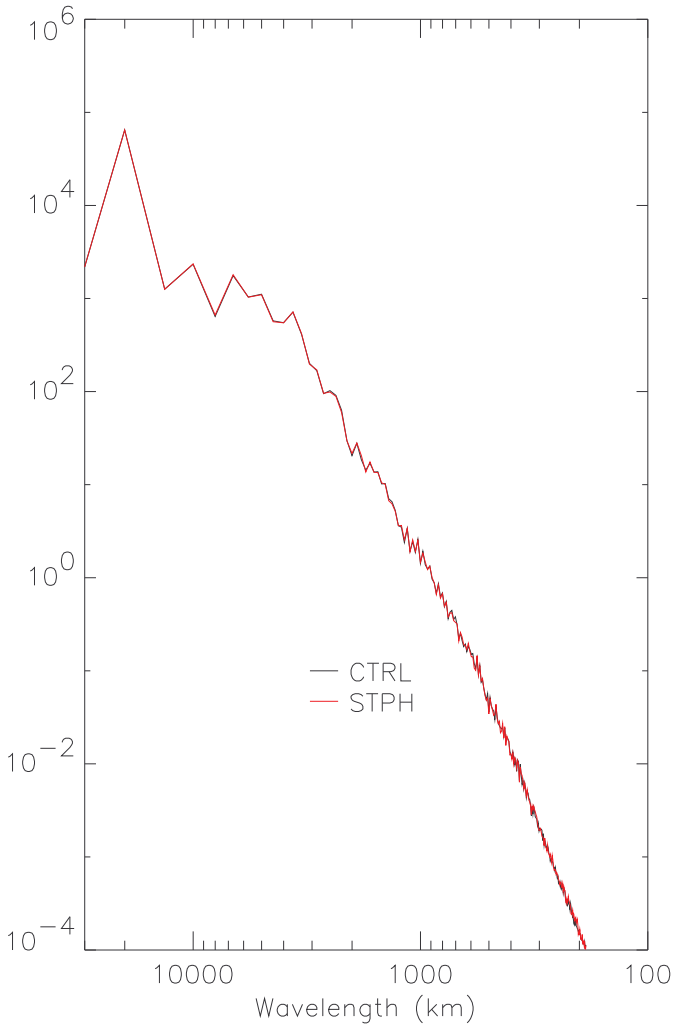

Figure 9. Kinetic energy spectra at 500 hPa and T+72 for the control simulation (black line) and one of the runs with the STPH scheme for the 20040326 case.

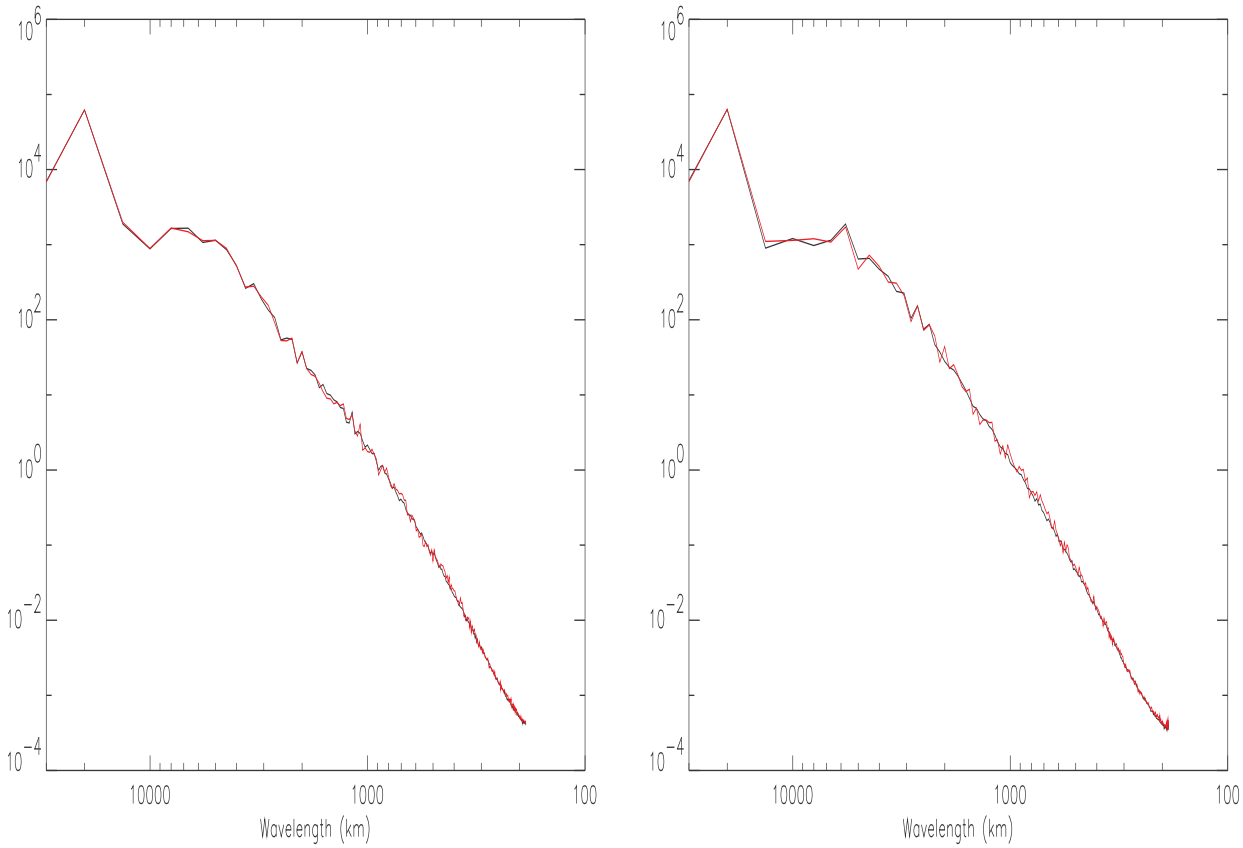

Figure 10. Kinetic energy spectra at 500 hPa for the IC (only perturbing the initial conditions) and ME (only including the STPH scheme as representation of the model error) ensembles at **(a)** T+12 and **(b)** T+72. The plotted energy spectra corresponds to the average from the 8 ensemble members.

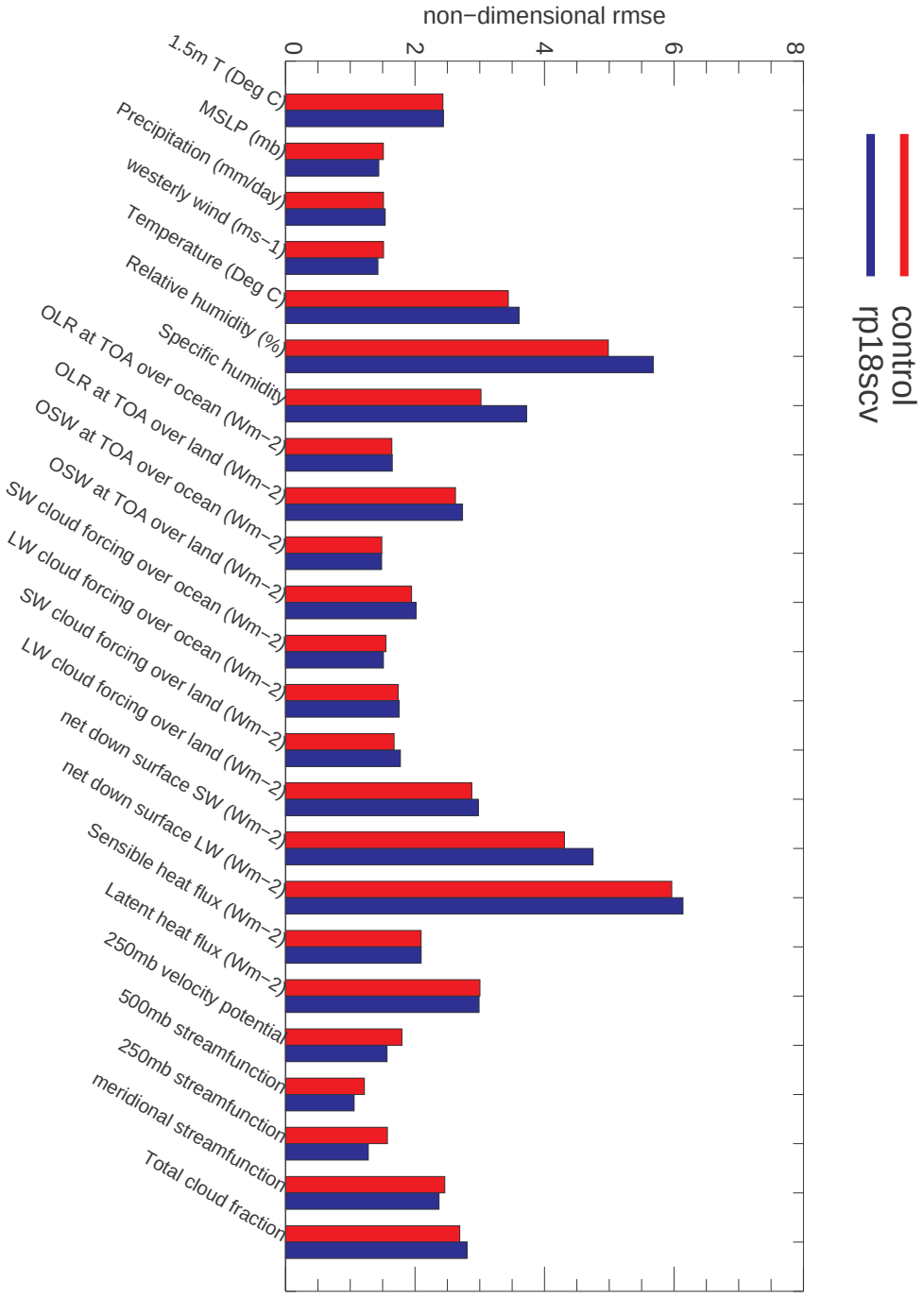

Figure 11. Values of the Climate Prediction Index (CPI) and its components. The 24 components represent surface and atmospheric variables, and are calculated as the r.m.s. difference between simulated and observed present-day climatological mean patterns divided by the r.m.s. value of the standard deviation of simulated interannual variations. The CPI is calculated as the r.m.s. value of the 24 components. All components are weighted equally, apart from the fields of cloud cover which receive a relative weight of 1/3 since observations of high, medium and low cloud are interdependent for a given thickness category. OLR, outgoing longwave radiation; OSW, outgoing shortwave radiation; LW, longwave radiation; SW, shortwave radiation; TOA, top of atmosphere; MSLP, mean sea level pressure.

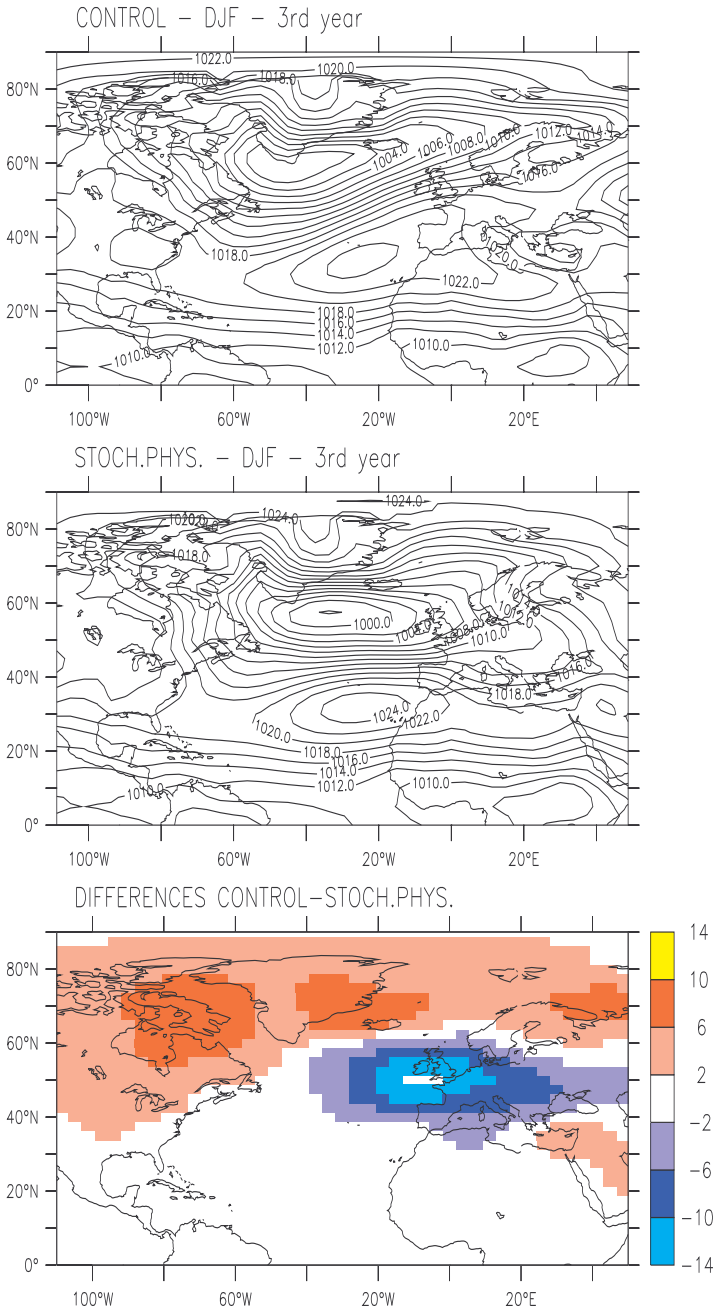

Figure 12. Mean sea level pressure for the winter period (December, January, February) at year 3. The control run is shown in the top panel, the run including the STPH scheme is shown in the central panel and the differences between both are shown in the bottom panel.

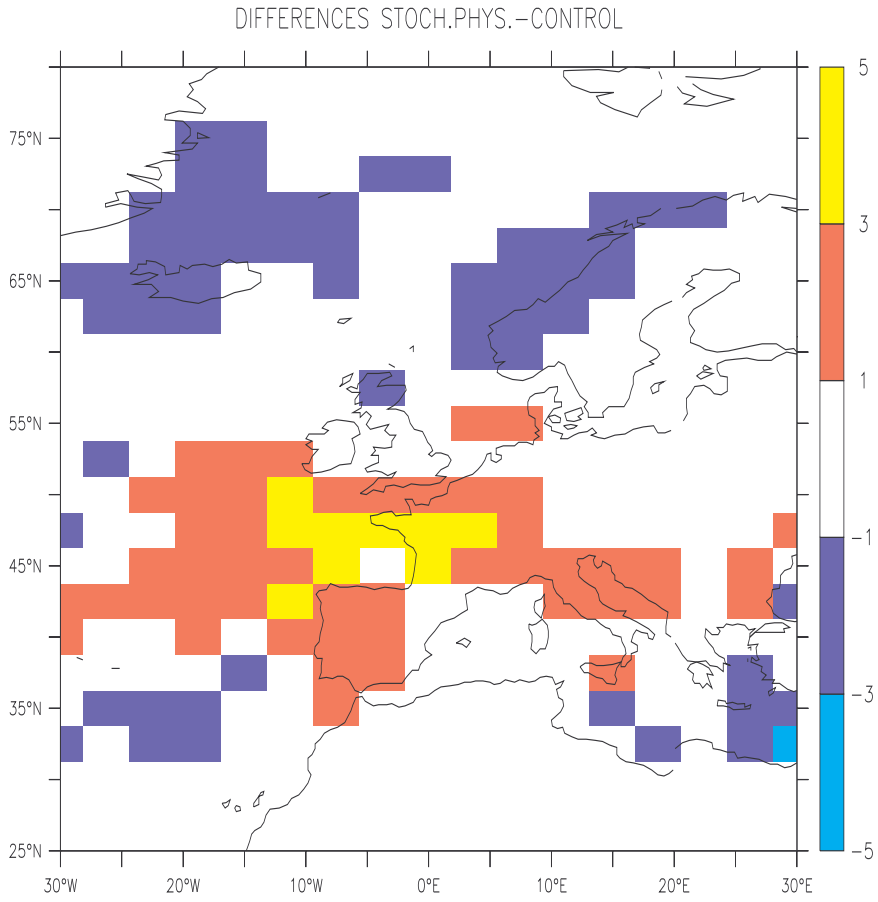

Figure 13. Differences in precipitation (mm/d) for the winter period (December, January, February) at year 3 between the control run and the run including the STPH scheme.